



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/036,577	12/31/2001	Andrew V. Anderson	5038-175	4052

7590 11/16/2004

STEVEN P. SKABRAT
BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP
12400 WILSHIRE BOULEVARD,
SEVENTH FLOOR
LOS ANGELES, CA 90025

EXAMINER

LERNER, MARTIN

ART UNIT PAPER NUMBER

2654

DATE MAILED: 11/16/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

10/036,577

Applicant(s)

ANDERSON ET AL.

Examiner

Martin Lerner

Art Unit

2654

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
 - If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
 - If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
 - Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☐ Responsive to communication(s) filed on ____.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1 to 29 is/are pending in the application.
- 4a) Of the above claim(s) ____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) ____ is/are allowed.
- 6) ☒ Claim(s) 1 to 29 is/are rejected.
- 7) ☐ Claim(s) ____ is/are objected to.
- 8) ☐ Claim(s) ____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☒ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on ____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
- Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. ____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date ____.
- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. ____.
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other: ____.

DETAILED ACTION

Specification

1. The abstract of the disclosure is objected to because it begins with a sentence fragment: "A speech recognition system." Correction is required. See MPEP § 608.01(b).

Claim Rejections - 35 USC § 112

2. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

3. Claims 13 and 14 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Claim 13 recites the limitation "the controller". There is insufficient antecedent basis for this limitation in the claim. Claim 13 should depend upon claim 12, which recites "the controller".

Claim Rejections - 35 USC § 102

4. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

Art Unit: 2654

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

5. Claims 1 to 6, 8 to 27, and 29 are rejected under 35 U.S.C. 102(b) as being anticipated by *Roberts et al.*

Regarding independent claim 1, *Roberts et al.* discloses a speech recognition system, comprising:

“at least one recognizer to produce output signals from audio input signals” – when step 111 detects an utterance (“audio input signals”), it causes the program to advance to step 119, which stores the token produced by step 118 in a memory buffer called TEMP_TOK; if the recognition mode has been set to TEXTMODE, step 121 causes step 123 to perform TEXTMODE recognition upon TEMP_TOK; TEXTMODE recognition is the normal recognition mode which enables the user to dictate words for inclusion in the textual output (“output signals”) of the system (column 8, lines 17 to 50: Figure 1: Steps 111, 121, and 123);

“a feedback module to generate feedback data” – if the recognition mode has been set to EDITMODE, step 120 causes step 122 to perform EDITMODE speech recognition on the token stored in TEMP_TOK; selection commands 125 for EDITMODE are “pick_one”, “pick_two”, etc., edit menu choice commands 126, such as “edit_one”, “edit_two”, etc., and letter commands 127, such as “starts_alpha”, “starts_bravo”, etc. (column 8, lines 17 to 50: Figure 1: Steps 120, 125, 126, and 127); commands for EDITMODE permit a user to provide “feedback” for correctness of speech recognition; selection commands 125, edit menu choice commands 126, and letter commands 127 are “feedback data” from a user.

Regarding independent claim 11, *Roberts et al.* discloses a speech recognition system, further comprising:

“wherein the speech recognizer is adapted to receive feedback data and adjust operation based upon the feedback data” – after step 178 stores a confirmed word in a language context buffer, step 180 uses the confirmed word to update the language model used by the recognition system (column 13, lines 44 to 60: Figure 1: Steps 178 and 180); updating a language model is equivalent to adapting a speech recognizer and adjusting its operation based upon confirmation (“feedback data”).

Regarding independent claims 16 and 25, *Roberts et al.* discloses a speech recognition method and machine-readable code, comprising:

“converting an audio input signal to an output signal” – when step 111 detects an utterance (“an audio input signal”), it causes the program to advance to step 119, which stores the token produced by step 118 in a memory buffer called TEMP_TOK; if the recognition mode has been set to TEXTMODE, step 121 causes step 123 to perform TEXTMODE recognition upon TEMP_TOK; TEXTMODE recognition is the normal recognition mode which enables the user to dictate words for inclusion in the textual output (“an output signal”) of the system (column 8, lines 17 to 50: Figure 1: Steps 111, 121, and 123);

“estimating a correctness measure wherein the correctness measure expresses if the output signal is a correct representation of the audio input signal” – a score (“a

Art Unit: 2654

correctness measure”) is computed for each time aligned match between the acoustic information in each frame and the acoustic model of the node against which it is time aligned; the words with the lowest sum of distance are then selected as the best scoring words (column 8, line 58 to column 9, line 7: Figure 3: Steps 129 to 132);

“forming a feedback data element wherein the element consists of at least one of the audio input signal, the output signal, and the correctness measure” – step 174 confirms the top choice, or best scoring word, from the recognition; step 176 displays the choices from the recognition of the token just saved, with the choices displayed in order, with the top choice, or best scoring word first, and with each choice having next to it a function key number “f1” through “f9” (column 12, lines 56 to 66: Figure 1: Steps 174 and 176); confirmation of word choices by a user provides a feedback data element through selection by function keys, where feedback involves at least scoring (“the correctness measure”) and confirmation of a word choice (“the output signal”).

Regarding claims 2, 3, 12, 13, 15, 21, and 26, *Roberts et al.* discloses a block diagram of a computer program for coordinating output of text by speech recognition (“production of the output signals”) and editing by selection commands 125, edit menu choice commands 126, and letter commands 127 (“adaptable to provide the feedback data to the recognizer”) (Figure 1); the computer program is “a controller”.

Regarding claim 4, *Roberts et al.* discloses using the confirmed word to update the language model used by the recognition system; for each pair of words W1, W2, the probability of W2 is updated by the number of counts for how often the pair occurs as

Art Unit: 2654

successive words in the text (column 13, lines 44 to 60: Figures 1 and 9); a language model of a probability of W2 given W1 is "a grammar file"; thus, updating a language model based upon confirming a word is equivalent to "modifying a grammar file based on the feedback data."

Regarding claims 5, 6, 17, and 27, *Roberts et al.* discloses storing confirmed words ("the feedback data") in SAV_TOK ("a storage"); step 214 finds all the tokens previously stored in the tokenstore in association with the just confirmed word and builds a new acoustic model ("speech models") for that word with those tokens; step 216 stores this acoustic word model with the other acoustic word models (column 15, line 58 to column 16, line 6: Figure 1: Steps 214 and 216); building a new acoustic model from a confirmed word is equivalent to "updating speech models based on the feedback data."

Regarding claim 8, *Roberts et al.* discloses TEXTMODE recognition produces recognized text; EDITMODE recognition produces command signals (column 8, lines 17 to 50: Figure 1).

Regarding claims 9 and 22, *Roberts et al.* discloses generating feedback based upon language model filtering so that words which the language model indicates are most probable in the current context are more likely to be selected (column 9, lines 1 to 7); a language model involves "grammar files" (column 13, lines 44 to 60: Figures 1 and 9); also, each of the displayed choices are "output signals".

Regarding claim 10, *Roberts et al.* discloses generating feedback based upon user choice editing by selection commands 125, edit menu choice commands 126, and

Art Unit: 2654

letter commands 127 (column 8, lines 37 to 50), or of function keys "f1" through "f9" (column 15, lines 27 to 40); these commands are "information received through an application programming interface".

Regarding claim 14, *Roberts et al.* discloses real time feedback as each word is recognized.

Regarding claims 18 and 29, *Roberts et al.* discloses tokens are saved only for confirmed words for adaptive speech recognition, i.e. a word that was confirmed as being correct (column 16, lines 7 to 22).

Regarding claim 19, *Roberts et al.* discloses language model filtering, where the score of a word depends upon a language model reflecting the probability of a word occurring in the present language context (column 9, lines 1 to 7).

Regarding claim 20, *Roberts et al.* discloses at least updating an acoustic model of a confirmed word ("updating acoustic models based on the feedback data") (column 15, line 58 to column 16, line 6).

Regarding claim 23, *Roberts et al.* discloses assigning a TEMP_TOK identifier to the token produced by an utterance for word confirmation ("as part of the feedback data element") (column 8, lines 17 to 21; Figures 1 and 2).

Regarding claim 24, *Roberts et al.* discloses confirmation of a word through language model filtering of a present language context ("identifying relevant contextual information") (column 9, lines 1 to 7).

Claim Rejections - 35 USC § 103

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

7. Claim 7 is rejected under 35 U.S.C. 103(a) as being unpatentable over *Roberts et al.* in view of *Thelen et al.*

Roberts et al. only discloses one speech recognizer, and omits multiple recognizers and a predictor to select a best performing recognizer from feedback data. However, *Thelen et al.* teaches speech recognition having parallel large vocabulary recognition engines 331, 332, 333, where a model selector 360 is used to select at least one of the speech recognizers in dependence on a recognition context. (Column 7, Line 30 to Column 8, Line 5: Figure 3) A stated advantage is to provide a recognition system that is better capable of dealing with huge vocabularies. (Column 1, Lines 53 to 55) It would have been obvious to one having ordinary skill in the art to provide multiple speech recognizers and a selector to select a best performing recognizer based upon a recognition context as taught by *Thelen et al.* in the speech recognition system of *Roberts et al.* for the purpose of providing a recognition system that is better capable of dealing with huge vocabularies.

Art Unit: 2654

8. Claim 28 is rejected under 35 U.S.C. 103(a) as being unpatentable over *Roberts et al.* in view of *Ortega*.

Roberts et al. discloses a speech recognition system providing updates and adaptation of acoustic models and language models for confirmed words. Thus, *Roberts et al.* does not expressly say that audio input signals are only stored for which the correction status indicates a correction was necessary. However, *Ortega* teaches deferred correction for speech recognition systems, where a file log identifies changes to a language model and any new words added through correction. Thus, there is an advantage that a speech file can be updated on another system. (Column 1, Line 44 to Column 2, Line 6) It would have been obvious to one having ordinary skill in the art to provide a log file only for words having a correction status indicating that correction was necessary as taught by *Ortega* in the speech recognition system of *Roberts et al.* for the purpose of permitting deferred correction on another system.

Conclusion

9. The prior art made of record and not relied upon is considered pertinent to Applicants' disclosure.

Woodward, Juang, Waibel et al., Martino et al., Maes, and Pérez-Méndez et al. disclose related art.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Martin Lerner whose telephone number is (703) 308-


Art Unit: 2654

9064. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (703) 305-9645. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

ML
11/9/04


Martin Lerner
Examiner
Group Art Unit 2654